WORK PLAN AND METHODOLOGY

Methodology

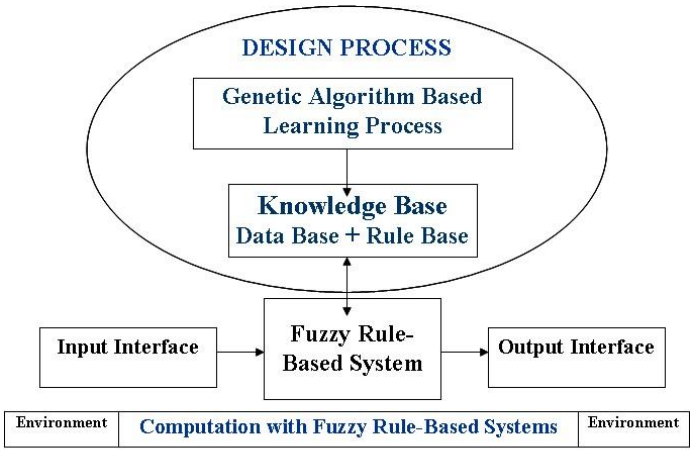
- Deep Literature Review on Fuzzy Sets and Fuzzy Logic
- Building association between fuzzy genetic approach and clustering for data mining and machine intelligence.
- The predictive approach mining for dynamic cluster formation for assorted datasets
- Implementation of the proposed approach in suitable simulation tool

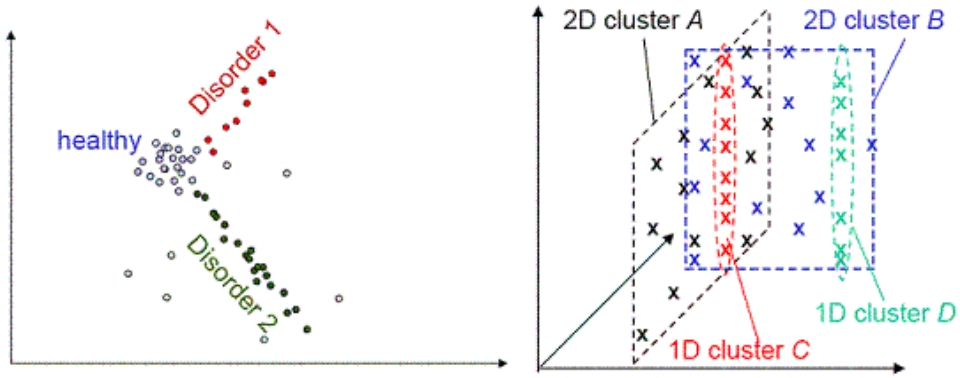
Clustering is a well-studied data mining problem that has found applications in many areas. For example, clustering can be applied to a document collection to reveal which documents are about the same topic. The objective in any clustering application is to minimize the inter-cluster similarities and maximize the intra-cluster similarities. There are different clustering algorithms each of which may or may not be suited to a particular application. The traditional clustering paradigm pertains to a single dataset. Recently, attention has been drawn to the problem of clustering multiple heterogeneous datasets where the datasets are related but may contain information about different types of objects and the attributes of the objects in the datasets may differ significantly. A clustering based on related but different object sets may reveal significant information that cannot be obtained by clustering a single dataset.

Problem Formulation

- There is need to develop a novel approach for dynamic clustering using metaheuristic algorithms
- The existing methods of clustering are not efficient as these approaches find out and group the data items in exact matching criteria
- There is need to design and implement the dynamic algorithmic approach that can cluster and display the data items so that there is non-biased clustering of data items in multiple domains.

- The fuzzy implementation is required in the cluster formation so that the exact matching can be reduced to the fuzzy based matching. Using this methodology, the data items or elements or inputs can be placed using many to many relationship.





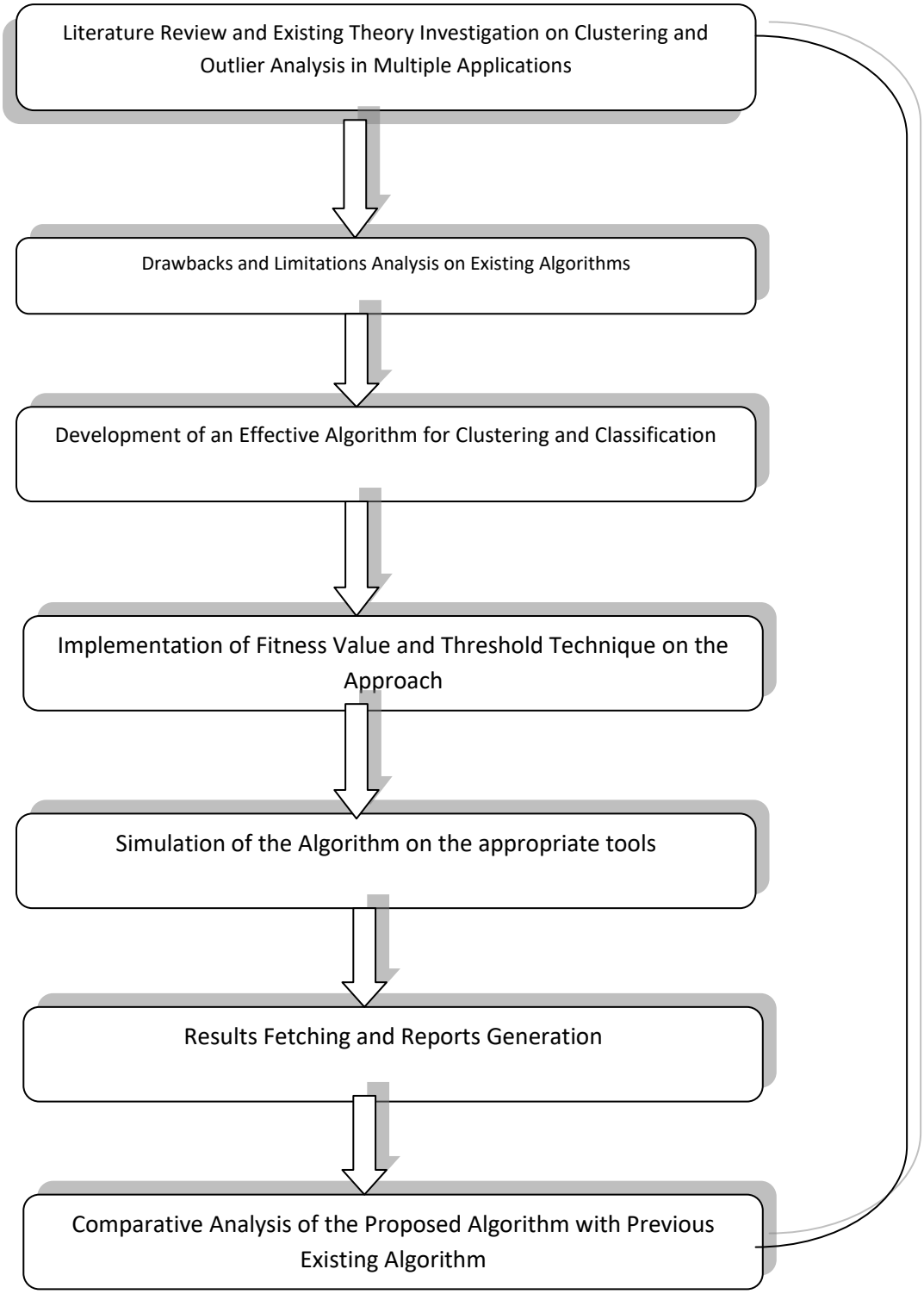

Literature Review and Existing Theory Investigation on Clustering and Outlier Analysis in Multiple Applications
Drawbacks and Limitations Analysis on Existing Algorithms
Development of an Effective Algorithm for Clustering and Classification
Implementation of Fitness Value and Threshold Technique on the Approach
Simulation of the Algorithm on the appropriate tools
Results Fetching and Reports Generation
Comparative Analysis of the Proposed Algorithm with Previous Existing Algorithm

Figure 4.1 – Flow Diagram

Work Plan

Phase 1: Deep Analytics from Literature Review and Extraction of the Novel Approaches in Practice

Phase 2: Dataset formation and Extraction of Feature Points

Phase 3: Extraction of results from Classical Approach and Proposed Approach

Phase 4: Analysis of the parameters from Classical and Proposed Approach in terms of multiple parameters

- Execution Time
- Turnaround Time
- Cost Factor
- Complexity
- Overall Efficiency and Performance
- Error Factors